DOCUMENT RESUME

ED 159 056 SE 024 909

AUTHOR Good, Thomas L.
TITLE The Missouri Mathematics Effectiveness Project: A Program of Naturalistic and Experimental Research.
PUB DATE Mar 78
NOTE 31p.: Paper presented at the annual meeting of the American Educational Research Association (Toronto, Canada, March 27-31, 1978)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS *Achievement; *Educational Research; *Effective Teaching; Elementary Education; *Elementary School Mathematics; *Elementary School Teachers; *Instruction; *Teacher Behavior; Teacher Education; Teacher Influence; Teaching Styles
IDENTIFIERS Research Reports

ABSTRACT

Teacher behavior associated with effective mathematics instruction in elementary schools had earlier been characterized and translated into an instructional program. Key instructional behaviors identified were: daily review, development, seatwork, homework assignment, and special reviews. A vigorous test of this program was undertaken using 40 fourth-grade classroom teachers. Teacher effectiveness was measured by student gains on scores on a standardized achievement test. The performance of the treatment group significantly exceeded the performance of the control group. A complete analysis is reported. (MP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED159056

SE

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Thomas L. Good

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

APR 7 10 55 AM '78

The Missouri Mathematics Effectiveness Project: A Program of Naturalistic and Experimental Research[1,2,3]

Thomas L. Good

University of Missouri--Columbia

1. This paper was prepared for the 1978 annual meeting of the American Educational Research Association, Toronto, Canada.

2. This research was partially supported by the National Institute of Education (Grant NIE-G-77-0003). The options expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement should be inferred. The author also acknowledges the support provided by the Center for Research in Social Behavior, University of Missouri, Columbia, Missouri, and the typing of the manuscript by Sherry Kilgore. The author also wants to acknowledge the colleagues who are presently working with him on the experimental phase of the research: Dr. Doug Grouws, Howard Ebmeier, Terrill Beckerman, Dianne Hunter, and Dr. Sharon Schneeberger.

3. The author would also like to acknowledge the support of participating teachers and principals in the Tulsa Public School System and to Dr. Jack Griffin, Associate Superintendent, and to Mr. J. W. Hosey, Mathematics Coordinator, for their consistent encouragement and support throughout the project.

SE 024 909

Background

In 1970, findings from teacher effectiveness studies were inconsistent and the level of accumulated knowledge in the field was low. Although various explanations for the lack of productivity have been offered (Dunkin and Biddle, 1974), perhaps the chief restraint on previous research was the use of purposeful teacher samples. Purposeful sampling procedures do not guarantee that important levels of behavioral (teachers teach differently) or product (students achieve more or less than predicted) variance will be included within the research design.

Since 1970 several different investigators have attempted to build product variance into their studies (i.e. identify teachers who obtain more than expected student gain) and have attempted to explain variance in achievement effects through intensive, naturalistic observation procedures. Rosenshine (1977) and Medley (1977) have reviewed most of this literature. It is beyond the scope of the present paper to repeat or even summarize those positions.

However, it is useful to note that within less than a decade, the field has moved from a state of confusion to a point at which experimental studies are being designed (upon a data base) with some optimism. The purpose of this paper is to review the efforts of one research program: from naturalistic observation to experimental intervention.

For the past several years, Dr. Doug Grouws and I have been engaged in a program of research attempting to characterize behaviors associated with effective mathematics instruction in elementary schools. Assisted by a grant from the National Institute of Education we began a program of naturalistic research in a large midwestern city five years ago.

The initial school district was chosen because a common textbook was used in all mathematics classrooms and because the composition of students in schools throughout the district was comparable. These characteristics were desirable because the investigators wanted to make inferences about desirable teacher behavior. By controlling for differences in curriculum materials and student home backgrounds, it would be possible to argue with greater confidence that observed differences in student achievement were due to teachers and not to context effects.

The second step was to identify a group of teachers who were consistent and relatively effective or ineffective in obtaining student achievement results. To estimate teachers' effectiveness, residual gain scores were computed for their students using their pre- and post-scores on a standardized achievement test (the student's pre-test score was used as a covariate in a linear model.

It was deemed important (given the expense of collecting and analyzing observational data) to select an observational sample of teachers who were consistent over consecutive years in their impact upon student achievement, and who were also notably different in their impact. That is, teachers were selected for observation who regularly obtained more (and teachers who obtained considerably less) achievement from students than did other teachers who taught similar students under similar circumstances.

Differences between More and Less Effective Teachers

After observing target teachers repeatedly, a behavioral profile was constructed for the relatively effective and ineffective teachers. Encouragingly, there emerged a set of factors that consistently separated the more and less effective teachers. These naturalistic findings were integrated with the recent research of others and translated into an instructional program (see Good and Grouws, 1977a).

It is beyond the scope of this paper to summarize, in detail, those teaching behaviors that differentiated more and less effective teachers or to summarize the entire instructional plan that was presented to teachers in the experimental stage of the research. These details can be found elsewhere (Good and Grouws, 1975; 1977a; 1977b).

The highlights of the instructional plan (that was based upon our earlier naturalistic research and the relevant work of others) are presented in Table 1. Most of the variables that appear in Table 1 are familiar ones and development is the only variable that would appear to need a definition and rationale.

Development: A Definition

The development portion of the mathematics period is that part of the lesson devoted to establishing comprehension of skills and concepts. It should be viewed as a continuum which runs from developing understanding to allowing for meaningful practice in a controlled setting. During all stages of the developmental portion of the lesson, both ends of the continuum may be present to some degree. However, in general, the comprehension emphasis with very little practice will come at the initial part of the lesson, then toward the middle of the lesson, practice with process feedback from the teacher will become quite prominent, and finally in the latter portion of the lesson there will be controlled practice with meaningful explanations given as necessary.

The role of the teacher in the first part of the lesson, the comprehension phase, is to use instructional strategies that help students to understand clearly the material being studied that day. In this portion, of the lesson, emphasis is placed upon comprehension rather than rote memorization. Activities which often focus on comprehension include

teacher explanations and demonstrations, and may include use of manipulative materials to demonstrate processes and ideas, use of concrete examples in order to abstract common features, making comparisons and searching for patterns, and class discussions.

During the middle portion of the lesson the number of questions posed to students may increase as the teacher begins to assess comprehension and provides them an opportunity to model processes already demonstrated and to verbalize the understanding they have developed. During this phase of the lesson the teacher may decide that further explanations and demonstrations are necessary or decide the controlled practice is appropriate since students seem to understand what they are doing.

In the controlled practice phase of the lesson the emphasis is on increasing proficiency; that is, increasing speed and accuracy. However, meaningful feedback is still given as necessary or requested.

Need for Development

Many problems arise in math classes where teachers give too little attention to development. Students exposed to such teaching frequently attempt to memorize rules for doing things and concentrate on mechanical skills. These rules have no meaning for the student (because developmental work was not done) and thus they are easily forgotten especially when new sets of rules are "learned." When students do not understand what they are doing, each new problem causes them great difficulty. Often the comment "We haven't done any of these before" is heard. When students learn without understanding, the ability to transfer skills to new situations is greatly reduced. Other negative results such as the inability to detect absurd answers and loss of self-confidence also occur. Thus, there are many compelling reasons to include a large measure of developmental work in mathematics lessons.

In the manual a definition and rationale are presented for each part of the lesson (similar to the material presented above for development) along with a detailed section on teaching practice. Here, teachers were given numerous examples of specific behaviors that they could use in the classroom to achieve project goals.

Space limitations prohibit a presentation of the definition, rationale, and teaching practice statement for each part of the lesson. However, it is useful to summarize the distinctive aspects of the treatment: (1) the program, in total, represents a system of instruction; (2) instructional activity is initiated and reviewed in the context of meaning; (3) students are prepared for each lesson stage so as to enhance involvement and to minimize student performance errors; (4) the principles of distributed and successful practice are built into the program; (5) active teaching is demanded.

The Study

Although certain of the instructional variables had been tested individually in other settings, the first experimental test of the entire instructional program commenced in the fall of 1977. Assisted by another grant from the National Institute of Education and the cooperation of the Tulsa Public School District it was possible to undertake a rigorous test of the program.

With the active assistance of administrators and building principals, it was possible to recruit a volunteer sample of 40 classroom teachers who used the semi-departmental organizational plan. The decision was made to do research within only one organizational pattern (although the principles apply, we feel, to different classroom organizational patterns, the need for different types of examples, etc., in the instructional manual

would have necessitated the construction and use of different manuals). The semi-departmental structure was chosen because it seemed to afford a basic classroom structure that was most comparable to the classroom structures in which the naturalistic research was originally conducted (e.g. no classrooms that were completely individualized). Choice of this structure also provided a rough control for instructional time as teachers did not keep the same students for the entire day. Hence, for most of the teachers there was pressure to end the mathematics class at a set time and reteaching later in the day was impossible.

On September 20, we met with all teachers and all school principals who had volunteered to participate in the project. Fourth grade teachers who taught using a semi-departmentalized[4] structure were invited to participate in the project and were told prior to the meeting that the research team had been developing a program designed to improve students' mathematics skills and encouraged to implement it in their classrooms. Eighty percent of the available population of teachers who were given information about the project volunteered for the program. As it turned out, most of the participants were teaching in low income Title I schools.

At the workshop, all teachers were reminded that the program was largely based upon an earlier observation of relatively effective and ineffective teachers. Teachers were also told that although we expected the program to work, the earlier research was correlational and that the present project was a test of those ideas. After a brief introduction, the 40 teachers (drawn from 27 schools) and their principals were divided into one of two groups: immediate feedback and delayed feedback.

---

[4]A semi-departmentalized structure calls for teachers to teach only two or three different subjects a day.

Schools were used as the unit for random assignment[5] to immediate and delayed feedback conditions. This was done to eliminate the difficulties that would doubtlessly follow by implementing the treatment in one class but not the other in the same school (e.g. why doesn't my child have homework?).

Teachers in the immediate feedback group were given an explanation of the program (the training lasted for 90 minutes). After the training session, treatment teachers were given the 45 page manual along with the instructions to read it and to begin to plan for implementation. In contrast, delayed feedback teachers (control) were told that they would not get the details of the program until February, 1978. Furthermore, they were told that it was hoped that this information might be especially useful because we would be able to supply them with individual information about their own classroom behavior as well as refined information about the program itself. Finally, delayed feedback teachers were told that their immediate role in the project was to continue to instruct in their own style. Efforts were made to create a strong Hawthorne group to maximize confidence in believing that any subsequent differences between immediate and delayed feedback groups would be due to program differences: not motivational differences. Given that control teachers knew that the research was designed to improve student achievement, that the school district was interested in the research, and that they were being observed, we feel reasonably confident that a strong Hawthorne control was created.

---

[5]Using information provided by the testing division in the school system, students were matched in terms of student SES populations and then randomly one school from the pair was assigned to the experimental condition. In the earlier naturalistic research, teachers were using only one textbook. In this project, two basic textbooks were in use, and it was impossible to match on both student characteristics and textbooks. Subsequent data analyses will focus on "text" and school effects.

Many educational improvement projects turn out to work only in the short run (when they are tried out by enthusiasts who make the program work not because the plan is good but simply because people work extra hard (e.g. plan more, etc.). For this reason, every effort was made to build in a strong "Hawthorne" control group. We hoped that the knowledge that the program was thought to be effective, designed to improve mathematics achievement, and the knowledge that they would be receiving information about their own classroom behavior would motivate control teachers to think more about their mathematics teaching. To the extent that this condition was created, it could be argued that differences in performance between control and treatment groups was due to the program: not to motivational variables.

However, at the other extreme, it was not intended to create so strong a "press" that teachers (because of concern) would seek out information from treatment teachers or would alter their instructional style trying to guess what the experimenters wanted. If control teachers changed their instructional behavior, then differences could be due to the fact that they were using a poorly thought out or inconsistent pattern of instruction.

Follow-Up Contact

Roughly two weeks after the treatment began (and after classroom observation was initiated for both groups of teachers), we returned to meet with treatment teachers (the immediate feedback group). The purpose of the 90 minute meeting was to respond to questions that teachers had about the meaning of certain teaching behaviors and to react to any difficulties that the teachers might have encountered.

The two major concerns expressed by teachers were with homework assignments (availability of materials and questions about what to do when students didn't hand in work) and the extent to which they could deviate from the treatment. We were able to help teachers with the supply problem (by having the district provide old, but comparable textbooks), but teachers were left largely on their own to solve the "completion" problem. A few suggestions were given for consideration, but no program definition was created for teachers to follow. Teachers were again told (as they had been at the first workshop and as written in the manual) ... "In general, we feel that the plan should be followed each day. However, we also realize that special circumstances will force you to modify the plan on occasion. Still, it is important that you follow the daily plan as frequently as you can."

Process and Product Measures

The treatment program started on October 3, 1977, and was terminated on January 25, 1978. During the course of the project all 40 teachers (with but few exceptions) were observed on six occasions. Observers collected information using both low and high inference process measures (see Good and Grouws, 1975, 1977b). Students were administered a standardized achievement test (SRA, Short Form E, blue level) in late September and in mid-December and a content test (a mathematics achievement test constructed by Professor Robert E. Reys at the University of Missouri) that attempted to measure the content that students had been exposed to during the program of research. Furthermore, an instrument measuring student learning styles and preferences and attitudes toward mathematics was administered in September and in January. At this time, the product and process assessment is but partially completed. The results on the

standardized test (which was administered in mid-December, 1977) have been analyzed reasonably well. The analysis of the mathematics content test (which was administered in mid-January, 1978) is but partially completed; however, the analysis work on the process and attitude data is just beginning.

Findings

Fortunately, at the debriefing session in February, control teachers consistently indicated that (a) they did think more about mathematics instruction this year than previously; (b) that they did not feel that they altered their behavior in any major way; (c) and that directly or indirectly they had not been exposed to program information. Hence, the Hawthorne control condition appeared to have been satisfactorily implemented.

Implementation

The second finding is that treatment teachers implemented the program reasonably well. If one is to argue that a program works (or is responsible for the change), it is important to show that teachers exhibited much more of the classroom behaviors related to the treatment than did control teachers. There were but two of the 20 treatment teachers who registered relatively low implementation scores. That is, those two teachers did not show as many of the program behaviors as we would have hoped. Development appears to be the only variable that teachers, as a group, had consistent trouble in implementing. However, there are still several variables related to development that have not been processed and inclusion of these variables (use of process feedback; ratio of process to product questions, etc.) may alter the interpretation somewhat. In general, teachers showed they could (and were willing to) implement the treatment after a minimal amount of training time. In summary, the treatment teachers were observed to be using the treatment program.

Given the complexity (several different behavioral requests involving sequences of behaviors) of the treatment it will be some time before it will be possible to provide a final, precise statement about the extent to which the treatment has been implemented. Presently, the behavioral process data collected in the project have not been analyzed. Thus, to estimate implementation we have been using a summary check list that observers filled out at the end of each observation. The information on the check list provides good but not total coverage of treatment behaviors. However, the additional information that is obtained subsequently is unlikely to change the basic finding that treatment teachers implemented the program reasonably well.

Using the check list, eight different implementation scores have been generated. Multiple definitions of implementation were used because it is possible to score program implementation in absolute and relative ways. For example, on the check list there are several questions related to development and seatwork. If one uses an absolute definition, all the questions about development are scored (even if the teacher did not include a development phase in the lesson). In the relative scoring scheme, if teachers did not engage in development on a particular observation they would be penalized only once, not repeatedly. Other scoring definitions were used on the assumption that several of the treatment behaviors were relatively minor and different definitions dropped out one to four of these variables and scored implementation in relative and absolute ways.

The rationale for the relative scoring was to control for frequency of occurrence. That is, did control teachers conduct development in the same way as treatment teachers but simply do it less often? By using relative scores it was possible to examine this question. As it turns out,

relative scoring systems did not reduce the gap between control and treatment teachers. That is, on all scoring definitions, treatment teachers appear to perform more of the treatment behaviors than did control teachers. Table 2 summarizes those behaviors included on the check list and which were used to estimate the extent to which teachers implemented the treatment. As can be seen in Table 3, all of the definitions significantly discriminated treatment and control teachers. Again, the reader is reminded that the implementation is not complete and several key qualitative (how stages of lessons were conducted) dimensions have not been analyzed. Subsequently, a complete implementation report will be available. This report will include qualitative variables as well as other variables (e.g. pacing) which have not been analyzed presently.

Impact of the Treatment on Student Performance

As can be seen in Table 4, the treatment group began the project with lower achievement scores than did the control group. The initial difference between experimental and control students was significant ($p<.001$).

These figures reveal that in two and one-half months of the project, the number of questions that were answered correctly by the average student in the experimental group increased from 11.75 to 19.94. Similarly, it can be seen that in terms of national norms, the percentile rank of the experimental group increased from a percentile of 25.7 to that of 57.3. Such results are truly impressive given the comparatively short duration of the project. Interestingly, the control group also shows a large gain, but their gains do not match those of the experimental group.

All experimental teachers taught mathematics to the class as a whole (as requested). However, 13 of the 20 control teachers taught mathematics to the whole class; whereas, the other seven taught mathematics to groups of students. Hence, pre and post differences for control whole class and control group teachers are presented separately in Table 4. As can be seen, control group teachers started

and ended the project with greater student achievement levels than did control whole class teachers and the experimental group. However, it should be noted that the raw gain of the experimental group was much higher than that of the control teachers who taught groups of students. Furthermore, in two and one-half months the experimental group virtually catches control teachers who used a group strategy.

Table 5 presents the results from an analysis of variance on post-test performance comparing experimental teachers with all control teachers (N=20) and with whole class control teachers (N=13). Regardless of which metric is used (grade level score, percentile score, or raw scores) the performance of the treatment group significantly exceeds the performance of the control group.

Table 6 presents the results from an analysis of variance on residual gain scores comparing the performance of experimental and control groups. Again, irrespective of the metric used, the performance of the treatment group significantly exceeds the performance of the control group. All of the residual means show a large positive discrepancy for the treatment group. That is, the experimental group showed considerably more achievement at the post-testing than was predicted by the pre-scores. In contrast, the control group showed a large negative discrepancy. Table 7 presents the initial results of an analysis of variance on the mathematics content test. In this table, the results from an analysis of variance on residual scores is presented (using raw scores on the SRA achievement test as the covariate). As can be seen in Table 7, the performance of the experimental group exceeds that of the control group.

Table 8 presents the correlations between implementation scores and residual gain performance on the standardized achievement test and the mathematics content test. In this table, the residuals are based upon the use of raw scores on the SRA test as the covariate and post performance on the SRA raw scores and the mathematics content test respectively. As can be seen in Table 8, all of the implementation definitions correlate positively and significantly with residual gain performance. However, it should be noted that the correlations are consistently higher between

implementation scores and performance on the standardized test than on the mathematics content test.

The lower correlations between implementation and the content test may be due to the procedures used in constructing the content test. The plan was to assemble a test that measured the content to which most students had been exposed. The test did not test the material that only a few teachers had reached. When the pacing data are analyzed (based upon teacher logs of the material covered daily) it will be possible to see if some teachers who had high residual gains on the standardized achievement test were penalized by a ceiling (content coverage) effect on the content test.

Table 9 presents the pre-means, post-means, and residual means (all based upon raw score units) on the SRA Standardized Achievement Test for each individual teacher. To facilitate interpretation, all of the experimental teachers appear in the top half of the table and all of the control teachers appear in the bottom half of the table (control teachers who teach groups of students are marked with an asterisk). Furthermore, teachers <u>within</u> the experimental and control conditions appear on the table in order of their residual mean rank.

Interestingly, within the control group (the naturalistic condition) it can be noted that teachers who use a whole class teaching strategy obtain both the best and worst results. Four of the five teachers who had the highest residual means taught the entire class; however, the lowest six teachers also taught with a whole class methodology. These results are a direct replication of our earlier naturalistic research (Good and Grouws, 1975; 1977b) in which it was also found that teachers who use a group teaching strategy fall in the middle of the distribution of residual gain scores.

Table 9 also illustrates the general effectiveness of the treatment as seven of the teachers with the highest top nine residual means were in the treatment group and none of the five lowest scores were in the treatment group. However, the impact of the treatment is not even across the treatment group: some teachers show considerably less gain than do other teachers. Hopefully, subsequent analysis of the classroom observation data will help to clarify the ways in which relatively more and less effective teachers achieved their results. However strong emphasis should be placed on the word relative for as can be seen in Table 9 all teachers' post-score means were higher than their pre-score means.

Finally, it should be noted that a sizeable, positive correlation (.60) was found between teachers' residual scores on the SRA test and the content test. Teachers who are high (or low) on one measure tend to be high (or low) on the other. Hence, in this particular setting the assessment coverage of the standardized achievement test appears to correspond reasonably well with the curriculum.

Summary

Much remains to be done with the data and more complete reports of the project will be available soon. Given the short period of the treatment program and the relative ease of implementation, the results appear striking. Importantly, teachers made the difference in student learning. It was not necessary to use expensive or difficult to prepare curriculum materials to achieve increases in student achievement. Also, these gains were made in urban, low income schools. That achievement increments can occur in such schools is aptly demonstrated by this project.

That control teachers and their students also showed marked improvement is probably due to the strong Hawthorne effect that was

purposefully built into the project (that is, knowledge that mathematics was the subject of interest; knowledge that teachers would receive information about their behavior during mathematics instruction; knowledge that the school district was interested in the project). Such Hawthorne motivation probably led control teachers to think more about their mathematical instruction and such proactive behavior (e.g. more planning) may have brought about increased achievement. However, the presence of a strong Hawthorne control makes it possible to argue that more confidence that the resultant differences between control and treatment classes are due to the instructional program: not because of motivational variables.

It is possible to argue that the large increases in student achievement are due to the decrement of mathematics skills over the summer. Hence, scores for all students were artifically depressed by the fall testing. However, this argument seems specious since the pre-test was administered six weeks after school had started. If the pre-test had been administered in the first two weeks, the argument might merit attention. Under the circumstances, the gains seem real (attributable to both program and motivational variables) and do not represent artifacts of measurement.

It is not suggested that the instructional program used in the study is the only or best approach to take for facilitating the mathematics achievement of students. However, it is argued that the instructional program is a good one for teachers who utilize and/or prefer a whole class organizational pattern for teaching mathematics in the middle elementary grades. Students, at this age, appear to benefit from direct instruction although it does not follow that all their mathematics instruction should be of this mode.

It should also be acknowledged that achievement in mathematics, although very important, is but one characteristic of effective teaching. All of us working on the project are interested in knowing more about the source of achievement gains. For example, did the experimental group do better than the control group because of gains in knowledge of mathematics facts and/or mathematics concepts and/or mathematical problem solving skills. Similarly, we have interest in knowing how the treatment influenced student attitudes toward mathematics classes.

Similarly, it will be important to examine the classroom behavior of teachers who are getting better gains and attempt to characterize the subtle factors that distinguish their behavior from that of other treatment teachers. Equally important is further consideration of the factors that inhibit or enhance student achievement in classroom organizations that call for the instruction of students in group settings. Increased conceptualization about instructional organization is needed if more insight into the teaching of mathematics is to be achieved.

Subsequently, when more of the project data have been analyzed, it will be possible to relate project findings to the findings produced by others and hopefully to articulate "integrative" concepts which can be tested in future research.

In closing I want to formally acknowledge the participating teachers for their involvement in the project. Both treatment and control teachers fulfilled their roles with excellent cooperation. The cooperation of teachers, principals, and district administrators was a key aspect in the success of the program. The present field experiment would appear to reflect the direct and immediate positive results than can be obtained when schools and researchers engage in collaborative efforts.

Table 1
Summary of Key Instructional Behaviors

Daily Review (First 8 minutes except Mondays)

a) review the concepts and skills associated with the homework

b) collect and deal with homework assignments

c) ask several mental computation exercises

Development (About 20 minutes)

a) briefly focus on prerequisite skills and concepts

b) focus on meaning and promoting student understanding by using lively explanations, demonstrations, process explanations, illustrations, etc.

c) assess student comprehension

1) using process/product questions (active interaction)

2) using controlled practice

d) repeat and elaborate on the meaning portion as necessary

Seatwork (About 15 minutes)

a) provide uninterrupted successful practice

b) momentum - keep the ball rolling - get everyone involved, then sustain involvement

c) alerting - let students know their work will be checked at end of period

d) accountability - check the students' work

Homework Assignment

a) assign on a regular basis at the end of each math class except Fridays

b) should involve about 15 minutes of work to be done at home

c) should include one or two review problems

Special Reviews

a) Weekly Review/Maintenance

1) conduct during the first 20 minutes each Monday

2) focus on skills and concepts covered during the previous week

b) Monthly Review/Maintenance

1) conduct every fourth Monday

2) focus on skills and concepts covered since the last monthly review

Table 2*

Comparisons of Treatment and Control Group Teachers on Selected Implementation Variables

		Treatment $\bar{x}$	Control $\bar{x}$	p-Value
1.	Did the teacher conduct review?	.91	.62	.0097
2.	Did development take place within review?	.51	.37	.16
3.	Did the teacher check homework?	.79	.20	.0001
4.	Did the teacher work on mental computation?	.69	.06	.001
5.	Did the teacher summarize previous day's materials?	.28	.25	.69
6.	There was a slow transition from review.	.07	.04	.52
7.	Did the teacher spend at least 5 minutes on development?	.45	.51	.52
8.	Were the students held accountable for controlled practice during the development phase?	.33	.20	.20
9.	Did the teacher use demonstrations during presentation?	.45	.46	.87
10.	Did the teacher conduct seatwork?	.80	.56	.004
11.	Did the teacher actively engage students in seatwork (first 1½ minutes)?	.71	.43	.0031
12.	Has the teacher available to provide immediate help to students during seatwork (next 5 minutes)?	.68	.47	.02
13.	Were students held accountable for seatwork at the end of seatwork phase?	.59	.31	.01
14.	Did seatwork directions take longer than one minute?	.18	.23	.43
15.	Did the teacher make homework assignments?	.66	.13	.001

*The means presented in the table represent percent of frequency of occurrence; time spent in each lesson stage will be presented later.

Table 3

Implementation Differences in Treatment and Control Teachers
Using Eight Different Scoring Definitions

Implementation Definitions	Treatment $\bar{x}$	Control $\bar{x}$	p-Value
I	7.2	4.3	.0001
II	7.7	4.4	.0001
III	3.9	2.4	.0001
IV	3.1	1.5	.0001
V	2.4	1.4	.0001
VI	3.2	2.3	.0034
VII	8.1	5.6	.0001
VIII	8.6	5.8	.0001

Table 4

Initial and Post Performance Status of Experimental and Control Students

I. All Treatment and All Control Teachers

	Pre Project Data			Post Project Data			Pre-Post Gain		
	Raw Score	Grade Equivalent	Percentile	Raw Score	Grade Equivalent	Percentile	Raw Score	Grade Equivalent	Percentile
Experimental	11.75	3.33	25.7	19.94	4.55	57.3	8.20	1.22	31.6
Control	12.61	3.45	30.0	17.60	4.21	48.5	4.99	.76	18.5

II. Control Whole Class Teachers and Control Group Teachers

	Pre Project Data			Post Project Data			Pre-Post Gain		
	Raw Score	Grade Equivalent	Percentile	Raw Score	Grade Equivalent	Percentile	Raw Score	Grade Equivalent	Percentile
Whole Class Control	11.43	3.26	24.29	15.91	3.95	42.1	4.48	.69	17.81
Group Control	14.3	3.74	36.0	20.24	4.63	58.5	5.94	.89	22.5

Table 5

Means and p-Values from an Analysis of Variance on the Raw Post-Scores of the SRA Achievement Test for Experimental and Control Teachers

I. Treatment vs. Control (Group Teachers Included

	Experimental	Control	p-Value
A. Grade level	4.55	4.21	.0001
B. Percentile	57.3	48.5	.0001
C. Raw	19.94	17.60	.0001

II. Treatment vs. Control (Group Teachers Not Included)

	Experimental	Control	p-Value
A. Grade level	4.55	3.95	.0001
B. Percentile	57.3	42.1	.0001
C. Raw	19.94	15.91	.0001

Table 6

Means and p-Values from an Analysis of Variance on Residual Gain Scores Comparing Treatment and Control Teachers on the SRA Test

I. Treatment vs. Control (Group Teachers Included)

	Treatment	Control	p-Value
Grade level scores	.24	- .21	.0001
Percentile scores	6.08	-5.29	.0001
Raw scores	1.74	-1.51	.0001

II. Treatment vs. Control (Group Teachers Not Included)

	Treatment	Control	p-Value
Grade level scores	.47	- .66	.0001
Percentile scores	5.64	-8.18	.0001
Raw scores	2.26	-3.27	.0001

Table 7

Means and p-Values from an Analysis of Variance on Residual Gain Scores Comparing Treatment and Control Teachers on the Mathematics Content Test

I. Treatment vs. Control (Group Teachers Included)

	Treatment	Control	p-Value
Content Test	.37	-.31	.08

II. Treatment vs. Control (Group Teachers Not Included)

	Treatment	Control	p-Value
Content Test	.48	-.68	.01

Table 8

Correlations between Implementation Definitions and Teachers' Residual Scores on the SRA Standardized Achievement Test and on the Content Test

Implementation Definition	Residual SRA Means	p-Value	Content Residual	p-Value
I	.44	.01	.33	.06
II	.46	.007	.32	.06
III	.46	.007	.38	.03
IV	.54	.001	.35	.05
V	.53	.001	.31	.08
VI	.38	.03	.33	.06
VII	.45	.008	.32	.07
VIII	.49	.004	.32	.07

Table 9

Pre, Post, and Residual Means on the Standardized Achievement Test by Individual Teacher

Experimental[1]	Pre SRA $\bar{x}$	Post SRA $\bar{x}$	Residual $\bar{x}$
1	13.05	25.95	6.62
1	10.00	22.54	5.47
1	13.49	23.58	4.24
1	9.89	21.11	4.11
1	16.05	23.56	3.67
1	13.32	23.43	3.47
1	17.83	26.78	3.13
1	14.48	23.12	2.76
1	12.09	21.70	2.74
1	14.70	23.36	2.13
1	11.46	20.35	1.93
1	11.00	20.33	1.60
1	11.81	19.39	1.13
1	7.62	14.75	-0.09
1	8.52	14.68	-0.57
1	8.63	15.22	-0.91
1	9.30	14.18	-1.54
1	19.00	23.13	-2.19
1	8.78	11.76	-3.76
1	8.08	11.50	-3.80

Table 9 (continued)

Control[2]	Pre SRA $\bar{x}$	Post SRA $\bar{x}$	Residual $\bar{x}$
2	18.70	28.04	3.61
*2	10.63	20.36	2.81
2	11.43	19.29	1.98
2	10.89	17.67	1.82
2	9.85	18.35	1.17
*2	15.24	21.57	0.50
*2	17.22	24.86	0.48
2	12.83	17.91	-0.76
*2	16.71	21.68	-0.87
2	10.30	15.50	-1.16
*2	14.63	19.94	-1.44
2	10.05	15.21	-1.93
*2	10.17	14.22	-2.52
*2	17.26	20.57	-2.60
2	13.12	16.05	-3.51
2	12.29	14.57	-4.11
2	11.33	13.60	-4.12
2	10.59	12.41	-4.77
2	7.74	9.20	-5.81
2	8.93	11.36	-5.95

1 = Experimental Group
2 = Control Group
*2 = Control teachers using a group format

References

Dunkin, M. and Biddle, B. The study of teaching. New York: Holt, Rinehart and Winston, 1974.

Good, T. and Grouws, D. Process-product relationship in fourth grade mathematics classrooms: Final Report of National Institute of Education (Grant NEG 00 3 0123). Columbia, Missouri: University of Missouri, 1975.

Good, T. and Grouws, D. Teaching Effects: A process-product study in fourth grade mathematics classrooms. Journal of Teacher Education, 1977b, 28, 49-54.

Good, T. and Grouws, D. Teaching manual: Missouri Mathematics Effectiveness Project. Report Series #132. Center for Research in Social Behavior, University of Missouri, Columbia, Missouri, 1977a.

Medley, D. Teacher competence and teacher effectiveness: A review of process-product research. American Association of Colleges for Teacher Education, Washington, D. C., 1977.

Rosenshine, B. New insights and questions from recent research in elementary schools. A paper read at the Annual Meeting of the American Educational Research Association, New York City, April 4-8, 1977.